Research article

Ruihuan Zhang, Yu He, Yong Zhang, Shaohua An, Qingming Zhu, Xingfeng Li and Yikai Su*

Ultracompact and low-power-consumption silicon thermo-optic switch for high-speed data

Abstract: Ultracompact and low-power-consumption optical switches are desired for high-performance telecommunication networks and data centers. Here, we demonstrate an on-chip power-efficient $2 \times 2$ thermo-optic switch unit by using a suspended photonic crystal nanobeam structure. A submilliwatt switching power of 0.15 mW is obtained with a tuning efficiency of 7.71 nm/mW in a compact footprint of 60 $\mu$m $\times$ 16 $\mu$m. The bandwidth of the switch is properly designed for a four-level pulse amplitude modulation signal with a 124 Gb/s raw data rate. To the best of our knowledge, the proposed switch is the most power-efficient resonator-based thermo-optic switch unit with the highest tuning efficiency and data ever reported.

Keywords: data transmission; optical switch; photonic crystal nanobeam; silicon photonics; tuning efficiency.

1 Introduction

Silicon photonic switches have attracted much attention thanks to their high compactness and the compatibility with the complementary metal-oxide-semiconductor fabrication process [1]. Typically, the switching between the on–off states of an on-chip silicon photonic switch can be achieved through thermo-optic (TO) tuning or electro-optic (EO) tuning process [2]. Benefitting from the large TO coefficient ($1.86 \times 10^{-4}$/K) of the silicon waveguide at 1.55 $\mu$m wavelength, silicon TO switches show advantages in high tuning efficiencies and compact footprints, therefore have been extensively studied for cost-effective and power-efficient optical switching networks [3–7].

Generally, most silicon TO switches are realized by applying overcladding heaters on top of the silicon waveguides, which can effectively prevent the metal absorption of the optical power [8–10]. However, these schemes are inherently power inefficient due to the large heat dissipations caused by the inevitable thermal spreading within the cladding oxides. In order to reduce the power consumption, it is feasible to decrease the heat dissipation, or to improve the efficiency in the tuning process of the switch. Various approaches and technologies have been proposed to reduce the power consumptions of the devices. In [11], a silicon Mach–Zehnder interferometer (MZI)–based switch was demonstrated using a highly $n$-type doped silicon heater that was directly contacted with the waveguide, thus decreasing the power dissipation in the heat conduction process. However, the fabrication process requires ion implantation and annealing. Some silicon TO switches added thermal isolation trenches by etching away the silicon substrates to suppress the heat diffusions, which helped to improve the tuning efficiencies [12–16]. However, the waveguides were still surrounded by the silica claddings, leading to limited performance improvements in the heat conduction processes. Other approaches are desired to further lower the power consumption of the silicon TO switch.

In a previous work [17], we demonstrated a power efficient switch by introducing high quality factor (Q-factor) resonators in the MZI switch. However, the accompanied narrow bandwidth with the high Q-factor limits the performance of the switch in high-speed data transmission. In this paper, we propose a scheme that can effectively lower the power consumptions of the on-chip TO switches. A suspended photonic crystal nanobeam (PCN)–based silicon TO MZI switch is fabricated as an example to illustrate the method. The PCN is designed with an ultrasmall mode volume ($\sim 0.243$ $\mu$m$^3$) and a reasonable Q-factor ($\sim 2000$) to enable a high tuning efficiency and a sufficient bandwidth for the switch.
Compared with the conventional MZI switches, the simulated tuning efficiency of the switch is dramatically improved by ~12 times. In the experiment, a submilliwatt switching power of 0.15 mW is obtained with a device footprint of 60 μm × 16 μm. The tuning efficiency of the switch is 7.71 nm/mW with a continuous tuning range wider than 25 nm. Moreover, a high-speed data switching experiment is carried out to study the switching performance. The results exhibit a low power penalty of less than 0.5 dB at a 124 Gb/s raw data rate with a four-level pulse amplitude modulation (PAM-4) format. A fast TO switching time of ~2 μs is also demonstrated. To the best of our knowledge, the proposed switch is the most power-efficient resonator-based TO switch unit with the highest tuning efficiency and capacity.

2 Optimization of the low power-consumption TO switch

The switching process of a conventional silicon TO MZI switch is realized by applying thermal energy to one of the waveguide arms to tune the optical phase [18]. During this process, the required phase shift of the switch can be decreased by introducing a resonant cavity into the waveguide, thus improving the tuning efficiency [19–22].

In a TO tuning microresonator, the resonance frequency shift caused by the applied thermal power can be expressed as [23–24]:

$$\Delta \omega = \frac{\omega_0}{2} \int \Delta \varepsilon(r) \cdot |E(r)|^2 \, dr,$$

where $\omega_0$ is the frequency of the Bloch mode, $\Delta \varepsilon$ is the variation of the cavity relative permittivity ($\varepsilon$) caused by the TO effect, $E(r)$ is the optical field distribution of the Bloch mode. The tuning power of the resonant cavity can be further obtained by considering the heat conduction process, which can be expressed as:

$$P = \frac{1}{k} \cdot \frac{C_{ws}}{Q_f} \cdot \frac{n}{m},$$

where $k$ is the TO coefficient of the silicon waveguide, $C_{ws}$ is the heat capacity of the heated waveguide, $n$ is the refractive index of the silicon waveguide, $Q_f$ is the Q-factor of the resonant cavity, $m$ is the overlap fraction between the optical field and the thermal region of the resonant cavity. The detailed mathematical description can be found in Supplementary material. According to Eq. (2), an optical switch with a low power consumption is supposed to employ a resonator with a small mode volume, a large $Q_f$ and a small $C_{ws}$.

Figure 1 shows the flow chart of the optimizing process of a low-power-consumption TO tuning resonator. Firstly, an ultrasmall mode volume is helpful to reduce the tuning power of a resonator by confining the optical field in a compact region. The tightly confined optical field in the cavity can effectively overlap with a given thermal region, thus increasing the value of overlap fraction ($m$) [24]. In addition, the small mode volume is also helpful to achieve a uniform thermal perturbation, which leads to an efficient tuning process. Secondly, as shown in Eq. (2), the resonator is supposed to have a high enough Q-factor ($Q_f$) to effectively decrease the power consumption. However, a high Q-factor will result in a limited bandwidth which impairs the high-speed data (e.g., 100 Gb/s or higher). Thus, the Q-factor of the resonator should be carefully tuned to enable the low power consumption of the switch while maintaining a sufficient bandwidth [25, 26].

Furthermore, removing the silica around the waveguide is also a viable method to reduce the power consumption by decreasing heat capacity ($C_{ws}$) [27]. Compared with a conventional silicon-on-insulator (SOI) waveguide, the suspended waveguide structure embedded in air exhibits a smaller heat capacity ($C_{ws}$) caused by the low heat conductivity coefficient of air, resulting in less thermal dissipation in the heat conduction process. Most of the
heating power is confined within the metal heater and the heated waveguide, thus a high tuning efficiency can be realized for the switch, which leads to a lower power consumption.

3 Switch structure and design principle

In order to verify the optimization method, we design a $2 \times 2$ silicon TO MZI switch by employing two suspended PCNs. The schematic configurations of the PCN and the switch are shown in Figure 2. In order to split or combine the transmission light power equally, two directional coupler-based 50:50 power splitters are employed at the input and output ports, respectively. The waveguide widths of the splitters are designed to be 450 nm. The coupling lengths and gaps are 10 μm and 200 nm, respectively. The lengths of the two arms of MZI are designed to be equal to ensure the constructive interference happens.

Two suspended PCNs with the same parameters are embedded in the two arms of MZI switch, and connected to the heaters by thin silicon slab waveguides. The schematic of the PCN cavity is illustrated in Figure 2(A). It consists of an array of 20 shallowly etched circular air holes to form a Fabry–Perot (F–P) cavity, which is symmetric with respect to its center. On each side of the cavity, the central six holes are adiabatically tapered to realize the phases matching between the waveguide mode and the Bloch mode, thus minimizing the cavity scattering loss of light. The tapered region is followed by four identical holes to serve as the reflector mirrors. The mode volume and Q-factor of the PCN cavity are iteratively optimized by using a 2.5D variational finite-difference-time-domain (FDTD) analysis tool. The optimization process can be found in our previous work [17]. In order to obtain a reasonable bandwidth for the switch, the Q-factor of the PCN is carefully tuned to be about 2000. The radii of the holes are optimized as: $r(n) = d(p + 15(n - 1))$, $(n = 1, \ldots, 6)$, $r(6) = r(6)$, $(n = 7, \ldots, 10)$, $d$ is the ratio of the radius to the period and is equal to 0.36, $p$ is the smallest period and is equal to 344 nm. As demonstrated in Figure 2(A), the radii are $r_1 = 124$ nm, $r_2 = 129$ nm, $r_3 = 134$ nm, $r_4 = 140$ nm, $r_5 = 145$ nm, $r_6 = 151$ nm, $r_7 = 151$ nm $(i = 6, 7, \ldots, 10)$, respectively. The periods are $p_1 = 344$ nm, $p_2 = 359$ nm, $p_3 = 374$ nm, $p_4 = 389$ nm, $p_5 = 404$ nm, $p_6 = 419$ nm $(i = 6, 7, 8, 9)$, respectively. In addition, the spacing $p_0$ between the two central holes is set to be 322 nm to introduce a defect for the resonant cavity. As shown in Figure 2(A), the optical field is confined within the cavity with a length of about 3 μm and a mode volume of about 0.243 μm$^3$. The width of the PCN waveguide is $w = 685$ nm to reduce the in-plane radiation loss of the cavity. Four 5-μm-long tapers are used to smoothly connect the 450-nm-wide waveguides with the 685-nm-wide PCN waveguides at the front and rear ports. Figure 3 shows the simulated transmission spectra of the device. The simulated insertion loss (IL) and cross talk values are 0.5 and $-30.7$ dB, respectively.

Furthermore, four air trenches are employed beside the two PCNs and heaters to prevent heat spreading in the slab waveguide. The thermal cross talk of the switch is negligible due to the low heat conductivity coefficient of air in the air trenches. The center to center distance of the two PCN arms is 15.4 μm, and the widths of the air trenches are 1.86 μm. In addition, the silica under the slab waveguide is removed, which prevents the heat diffusions in
the silica and silicon substrates. We have simulated the thermal distributions of the PCN waveguides using 3D finite element method simulation as shown in Figure 4. Compared with the conventional waveguide structure, the thermal field of the suspended structure is mostly confined within the PCN, benefiting from the air trench and the undercutting structure. The temperature of the suspended PCN increases from 293 to 668 K with a 1 mW heating power, which is about 12 times higher than the conventional PCN waveguide (from 293 to 324 K).

Considering the silicon TO effect and the electromagnetic perturbation theory, the greater temperature change indicates a larger increment of the refractive index, leading to a larger wavelength shift of the PCN. A simulated tuning efficiency of 15.62 nm/mW is obtained for a dual PCNs-based TO MZI switch. The detailed mathematical description of the TO tuning efficiency can be found in Supplementary material.

4 Device fabrication

The switch was fabricated on a SOI wafer, with a 220-nm-thick top silicon layer and a 3-μm-thick buried silica layer. The waveguides and PCNs were shallowly etched with a depth of 170 nm, while the air trenches were fully etched in the top silicon layer. The device was fabricated by using E-beam lithography technology (Vistec EBPG 5200+) and inductively coupled plasma etching process (SPTS DRIE-I). Two 300-nm-thick platinum heaters were sputtered beside the PCNs. Three pads were evaporated with 300-nm-thick gold and connected to the heaters as electrodes. The heaters and metal electrodes were fabricated by lift-off process with the length of 20 and 150 μm, respectively. Finally, the silica layer beneath the PCNs was etched by a dilute hydrofluoric acid solution to form the free-standing structure. The detailed fabrication processes of the suspended waveguide structure can be found in [28]. Figure 5 shows the microscope image and scanning electron microscope (SEM) image of the fabricated device structure. The footprint of the device is 60 μm x 16 μm. The red-dotted areas in the microscope photo and the SEM image are the air trenches, the white-boxed areas are the PCNs, and the green-dotted areas in the microscope photo are the 50:50 splitters.

5 Experimental results and discussion

The performance of the fabricated device was characterized by using a tunable laser (Keysight 81960), an optical power meter and two voltage-current source-meters (Keithley 2400). Four photonic crystal structure-based grating couplers were adopted at the two input ports and
two output ports to couple the light in and out of the device. The coupling loss of the grating coupler is 11.1 dB/facet at the central wavelength of 1550 nm according to the transmission result of a reference waveguide. The performances of the grating couplers are moderate which might be affected by the fabrication errors.

Figure 6(A) shows the unaligned transmission spectra measured at the drop port (Drop) and the through port (Thru), respectively. There is a resonance wavelength misalignment (\(-9.5 \text{ nm}\)) between the two PCNs, which is caused by the parameter deviations introduced during the fabrication process. We apply an electric power of 0.67 mW to tune the shorter resonance wavelength to make sure that the two PCNs operate in the same state. When the tuning power increases to 0.15 mW, the device is switched to the bar state. Further wavelength shifts are measured along with the tuning powers of 0.49, 1.15, and 1.78 mW. Note that the transmission spectra of the device are not flat, the ILs increase during the thermal tuning process, which may be attributed to the change of cavity losses induced by (1) the nonideal 50:50 splitters with limited operation bandwidths; and (2) nonuniform refractive index distribution induced by the heating power. These two factors cause performance deviations from the optimal designed values. In Figure 7(B), we use a curve to linearly fit the wavelength shift of the designed switch as a function of the tuning power. The slope of the fitting solid line corresponds to the tuning efficiency of the device, which is 7.71 nm/mW.

Furthermore, an on--off switching test is carried out to measure the response time of the switch. As shown in Figure 8, two 10-KHz square wave electric signals are loaded onto the heaters simultaneously. Then, the output signals are obtained at the drop port and through port, respectively. The signals show a rising time constant of \(-1.2 \mu\text{s}\) and a falling time constant of \(2.3 \mu\text{s}\) at the drop port, a rising time constant of \(2.2 \mu\text{s}\) and a falling time constant of \(1.0 \mu\text{s}\) at the through port.
We provide a brief summary for previously reported on-chip $2 \times 2$ TO switching units in Table 1. Resonator-based switches inherently exhibit more compact device footprints relative to nonresonant structures. Overall, the proposed switch shows balanced performances, and demonstrates the lowest switching power, the highest tuning efficiency, and the compact footprint in resonator-based $2 \times 2$ TO switches. While most previous characterizations were limited to the device performances only, we provide in this paper the bit error ratio (BER) measurement of such a switch with a 124 Gb/s PAM-4 data, as detailed below. This is the highest reported rate in a $2 \times 2$ TO switch, to the best of our knowledge.

Here we detail the transmission performance of a 62-GBaud PAM-4 signal through the PCN-based MZI switch for both the Bar and Cross states. The experimental setup and the transceiver digital signal processing (DSP) flow charts are shown in Figure 9. At the transmitter, a PAM-4 data stream is up-sampled by two times and then...
processed by a root raised cosine filter with a roll-off factor of 0.01. After resampling, the PAM-4 signal is sent to the 65-GSa/s arbitrary waveform generator (Keysight M8195A). The output signal of the AWG is amplified by an electrical amplifier and then drives a 25-GHz intensity modulator (IM) biased at the quadrature point of its transmission curve. A continuous-wave light from a distributed feedback laser is injected into the IM. After the electrical-to-optical conversion, the optical PAM-4 signal is boosted by an erbium-doped fiber amplifier (EDFA), followed by an optical band-pass filter (OBPF) to suppress the amplified spontaneous emission noise. After the OBPF, a polarization controller is used to adjust the polarization state of the PAM signal before entering the silicon chip. At the output port of the silicon chip, a variable optical attenuator is inserted to adjust the received optical power.

The receiver consists of an EDFA for preamplification, an OBPF, a 50-GHz photodetector (PD), and an 80-GSa/s digital storage oscilloscope (LeCroy 36Zi-A). In the receiver DSP, the signal is first resampled to a sampling rate of two samples per symbol. After the synchronization and matched filtering, a 201-tap linear feedforward equalizer (FFE) is applied. The equalizer coefficients are extracted.

**Figure 8:** Temporal responses of the switch. (A) The square-wave driving signal loaded onto the heater of PCN1. (B) The square-wave driving signal loaded onto the heater of PCN2. (C) The output signal from the drop port. (D) The output signal from the through port of the switch.

**Figure 9:** (A) Experimental setup and (B) digital signal processing (DSP) flow charts for the high data rate switching of the device. PM, power meter.

<table>
<thead>
<tr>
<th>Type</th>
<th>Device footprint (μm²)</th>
<th>Switching power (mW)</th>
<th>Tuning efficiency (nm/mW)</th>
<th>IL (dB)</th>
<th>Cross talk (dB)</th>
<th>Bandwidth (nm)</th>
<th>Data rate (Gb/s)</th>
<th>Modulation format</th>
<th>Switching time (μs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adiabatic bend-based MZI [11]</td>
<td>~5000</td>
<td>12.7</td>
<td>—</td>
<td>0.5</td>
<td>~20</td>
<td>70</td>
<td>—</td>
<td>—</td>
<td>2.2</td>
</tr>
<tr>
<td>Si-PCM [31]</td>
<td>~3 mm-length</td>
<td>1.41</td>
<td>—</td>
<td>2</td>
<td>~10</td>
<td>110</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Dual MRRs [12]</td>
<td>&gt;51 × 17</td>
<td>6</td>
<td>0.91</td>
<td>4</td>
<td>~20</td>
<td>0.2</td>
<td>12.5</td>
<td>OOK</td>
<td>9</td>
</tr>
<tr>
<td>MRR [32]</td>
<td>15 × 15</td>
<td>~3</td>
<td>0.39</td>
<td>~2</td>
<td>~9.5</td>
<td>0.32</td>
<td>1</td>
<td>—</td>
<td>2.91</td>
</tr>
<tr>
<td>Folded-waveguide MI [15]</td>
<td>400 × 130</td>
<td>0.05</td>
<td>—</td>
<td>3.3</td>
<td>~26</td>
<td>&gt;10</td>
<td>—</td>
<td>—</td>
<td>780</td>
</tr>
<tr>
<td>Photonic crystal [33]</td>
<td>15 × 15</td>
<td>3</td>
<td>1.17</td>
<td>&lt;1</td>
<td>~20</td>
<td>~3</td>
<td>—</td>
<td>—</td>
<td>2.91</td>
</tr>
<tr>
<td>Cascaded PCNs-based MZI [19]</td>
<td>150 × 30</td>
<td>0.16</td>
<td>1.23</td>
<td>1.5</td>
<td>~15</td>
<td>0.09</td>
<td>—</td>
<td>—</td>
<td>4.5</td>
</tr>
<tr>
<td>Dual PCNs-based MZI [17]</td>
<td>75 × 15</td>
<td>0.24</td>
<td>0.66</td>
<td>1.5</td>
<td>~15</td>
<td>0.12</td>
<td>—</td>
<td>—</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>This work</strong></td>
<td><strong>60 × 16</strong></td>
<td><strong>0.15</strong></td>
<td><strong>7.71</strong></td>
<td><strong>4.4</strong></td>
<td><strong>~13.5</strong></td>
<td><strong>0.68</strong></td>
<td><strong>124</strong></td>
<td><strong>PAM-4</strong></td>
<td><strong>2.3</strong></td>
</tr>
</tbody>
</table>

PCM, Phase change material; MI, Michelson interferometer.
from the training sequence using recursive least square (RLS) algorithm. To minimize the influence of the noise enhancement effect of the FFE, a 2-tap poster filter and the maximum-likelihood sequence decision are implemented. Finally, the PAM-4 demapping and BER calculation are performed.

Figure 10(A) shows the BER performance of the 62-GBaud PAM-4 signal in the presence of optical switching. The 7% hard-decision forward error correction threshold of $3.8 \times 10^{-3}$ is achieved for all the switching configurations ($I_1-O_1$, $I_1-O_2$, $I_2-O_1$, $I_2-O_2$). We also measured the BER curves for the worst switching state ($I_1-O_2$) as shown in Figure 10(B). Compared to the optical back-to-back (OBTTB) sensitivity of $-18.5$ dBm, a low penalty of $0.5$ dB is observed for the switching signal, indicating the bandwidth of the proposed switch is wide enough for the high-capacity switching at 124 Gb/s rate. The eye diagrams of the recovered PAM-4 signals at different switching configurations are provided in Figure 10(C).

**6 Summary**

In summary, we proposed an optimization method for the designing of a TO switch with a detailed mathematical derivation of the TO tuning process. A suspended PCN-based silicon TO MZI switch is designed to verify the feasibility of the method. Both the simulation and experimental results show significant improvement of the device performance compared with the conventional counterparts. The device is fabricated on a SOI platform with a small footprint of $60 \mu m \times 16 \mu m$. By thermally tuning the refractive index of the silicon PCN waveguides, the cross and bar states of the device can be switched. Only about $0.15$ mW tuning power is required for the cross-to-bar switching. Both ILs and cross talk values of the cross and bar state are measured as IL (bar) = $2.0$ dB, IL (cross) = $4.4$ dB, cross talk (bar) = $-13.8$ dB, cross talk (cross) = $-13.5$ dB, at the resonance wavelength of 1560.90 nm. The thermal tuning efficiency of the switch is $7.71$ nm/mW. The FWHM is measured as $0.68$ nm (i.e. $85$ GHz) at the through port with a calculated Q-factor of 2300. Moreover, system performance of high-speed data switching is also studied, which exhibits a low power penalty of less than $0.5$ dB at a 124 Gb/s raw data rate with the PAM-4 format.

**Acknowledgments:** The authors acknowledge the support of device fabrication by the Center for Advanced Electronic Materials and Devices (AEMD) of Shanghai Jiao Tong University (SJTU).

**Author contribution:** All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

**Research funding:** National Natural Science Foundation of China (NSFC) (Grant No. 61860206001/61835008).

**Conflict of interest statement:** The authors declare no conflicts of interest regarding this article.

**References**


Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/nanoph-2020-0496).